

Q95957 Substitute Sequence Listing 0409.TXT
SEQUENCE LISTING

<110> GAO, Guangxia
LIU, Shufeng

<120> The Recombinant Murine Leukemia Virus Reverse Transcriptase,
the Genes Encoding and the Method for Expressing it

<130> Q95957

<140> US 10/585,863

<141> 2006-07-11

<150> PCT/CN04/00039

<151> 2004-01-13

<160> 9

<210> 1

<211> 7488

<212> DNA

<213> artificial sequence

<220>

<223> plasmid DNA

<400> 1

cgcacccat	cgaatgtgtc	aaaaccttcc	gcggtatggc	atgatagcgc	ccggaagaga	60
gtcaattcag	gggtgggtgaat	gtgaaaccag	taacgtttata	cgatgtcgca	gagtatgccg	120
gtgtctctta	tcagaccggt	tccgcgctgg	tgaaccaggc	cagccacggt	tctgcgaaaa	180
cgcgggaaaa	agtggaagcg	gcgatggcgg	agctgaatta	cattcccaac	cgctgggacg	240
aacaactggc	gggcacaaacg	tcgttgctga	ttggcgcttc	cacctccagt	ctggccctgc	300
acgcgcgcgtc	gcaaatgtct	gcggcgatta	aatctcgcgc	cgatcaactg	gggtgccagc	360
tgggtgtgtgc	gatggtagaa	cgaagcgccg	tcgaagcctg	taaaagcgcg	gtgcacaatc	420
ttctcgcgca	acgcgtcagt	gggctgatca	ttaaactatc	gctggatgac	caggatgcca	480
ttgctgtgga	agctgcctgc	actaatgttc	cggcgcttatt	tcttgatgtc	tctgaccaga	540
cacccatcaa	cagtattatt	ttctcccatg	aagacggtac	gcgactgggc	gtggagcatc	600
tggctcgatt	gggtcaccag	caaatcgcgc	tgttagcggg	cccattaagt	tctgtctcgg	660
cgctctcgcg	tctggctggc	tggcataaat	atctcactcg	caatcaaat	cagccgatag	720
cggaaacggga	aggcgactgg	agtgccatgt	ccggttttca	acaaaccatg	caaatgctga	780
atgagggcat	cgttcccact	gcgatgctgg	ttgccaaaga	tcagatggcg	ctgggcgcaa	840
tgccgcgcat	taccgagttc	gggctgcgcg	ttggtgcgga	tatctcggta	gtgggatacg	900
acgataccga	agacagctca	gtttatatcc	cgccgttaac	caccatcaaa	caggattttc	960
gcctgtctggg	gcaaacccagc	gtggaccgct	gtgtgcaact	aaccaccctg	cgccccaata	1020
agggcacaatc	gctgtgtccc	gtctcactgg	tgaaaagaaa	gcgcgggagc	cgacagggtt	1080
cgcacaaacgc	ctctccccgc	cgcttggccg	attcattaat	gcagctggca	cataacggtt	1140
cccgaactgga	aagcgggcag	tgagcgaac	gcaattaatg	tgagttagct	cactcattag	1200
gcacaatttct	catgtttgac	agcttatcat	gcactgcacg	gtgcaccaat	gcttctggcg	1260
tcaggcagcc	atcgggaagt	gtggtatggc	tgctcaggtc	gtaaatcact	gcataattcg	1320
tgctgctcaa	ggcgcacttc	cgttctggat	aatgtttttt	gcgcgcacat	cataacggtt	1380
cttggcaata	tcttgaaatt	agctgttgac	aattaatcat	cggtctgcat	aatgtgtgga	1440
attgtgcagc	gataacaatt	tgaattctta	agattttgtga	ggggataaaca	gttcacacaa	1500
ggaaacagaa	tatgacctta	aaatagaag	atgagcatcg	gctacatgag	acctcaaaag	1560
agccagatgt	ttctctaggg	tccacatggc	tgcttgattt	ttctcaggag	caggcggaag	1620
ccgggggcat	gggactggca	gttgcgcaag	ctctctctgat	cataactctg	aaagcaacct	1680
ctacaccctgt	gtccataaaa	caatacccca	tgtcacaaga	agccagactg	gggatcaagc	1740
ccccataaca	gagactgttg	gagcgggaa	tcttggtacc	ctgccagctc	ccctggaaca	1800
cgcccctgtg	accggttaag	aaaccaggga	ctaagtatta	taggcctgtc	gagatcttga	1860
gagaagtcaa	caagcgggtg	gaagacatcc	accccacctg	gcccaacctc	tacacacctt	1920
tgagcgggtg	cccacgtccc	caccagtggg	acactgtgct	tgattttaag	gactgccttt	1980
tctgctctgag	actccaccct	accagtcagc	ctctcttcgc	ctttgagtgg	agagatccag	2040
agatgggaat	ctcaggacaa	ttgacctgga	ccagagctgc	acagggtttc	aaaaacagtc	2100
ccacctctgt	tgatgaggca	ctgcacagag	acctagcaga	cttcgggacg	cagcaccagg	2160

Q95957 Substitute Sequence Listing 0409.TXT

acttgatcct	gctacagtac	gtggatgact	tactgtctggc	cgccacttct	gagctagact	2220
gccaaacaagg	tactcggggc	ctgtttacaaa	ccctagggaa	cctcggggtat	cgggcctcgg	2280
ccaagaaagc	ccaattttgc	cagaaaacagg	tcaagtatct	gggggtatctt	ctaaaagagc	2340
gtcagagatg	gctcagctgag	gccagaaaag	agactgtgat	ggggcagcctt	actccgaaga	2400
cccttcgaca	actaaggggag	ttcctagggg	cggcaggctt	ctgtcgccct	tggatccctg	2460
ggttttcgaga	aatggcagcc	cccttgtacc	ctctcaccaa	aagcggggact	ctgttttaatt	2520
ggggcccgaga	caacaacaaag	gcctatcaag	aaatcaagca	agctcttcta	actgccccag	2580
ccctgggggtt	gccagattttg	actaagccct	ttgaactctt	tgctcgacgag	aagcagggct	2640
acgccaagag	tgctcattttg	caaaaaactgg	gacctttggcg	tcggcgcggtg	gcttacctgt	2700
ccaaaagagc	agaccacagta	gcagctgggt	ggcccccttg	cctacggatg	gtagcagcca	2760
ttgcgctact	gacaaaaggat	gcaggcaagc	taaccatggg	acagcagcta	gtcattctgt	2820
cccccatgc	agtagaggca	ctagtcaaac	aacccccga	ccgctggctt	tccaacgccc	2880
ggatgactca	ctatcaggcc	ttgcttttgg	acacggaccg	gggtccagttc	ggaccgggtg	2940
tagccctgaga	ccggcttacg	ctgtctccac	tgccctgagga	agggtgcaa	cacaactgcc	3000
ttgatattct	ggccgaagcc	cacggaaccc	gacccgacct	aacggaccag	ccgctcccag	3060
acgccgacaa	cactctgtac	acgaatggaa	gcagtctctt	acaagagctg	cagcgtaagg	3120
cgggagctgc	ggtagaccac	gagaccgagg	taattctgggc	taaaagccct	cgaccctgga	3180
catccgctca	gcgggctgaa	ctgatagcac	tcaccaggcg	cctaagaagt	gcagaaaggt	3240
agaagctaaaa	tgtttatact	gatagcgctt	atgcttttgg	tactgcccat	atccatggag	3300
aaataatacag	aaggcgtggg	ttgtctacat	cagaaggcaa	agagatcaaa	aataaagagc	3360
agatctttggc	ctactataaaa	gccctctttc	gtcccaaaaag	acttagcata	gcttactgtc	3420
caggacatca	aaagggaacc	agcgcggagg	ctagaggcaa	ccggatggct	caccaacgcy	3480
cccgaaggc	agcactcaca	gagactccag	acacctctac	cctctctcat	gacctcacc	3540
atcacatgag	tagagtgcag	ctgcaggcaa	gcttggcact	ggccgtcggt	ttacaacgct	3600
gtgactggga	aaaccctggg	gttaccacac	ttaatcgctt	tgccagacat	ccccctttcg	3660
ccagctggcg	taatacgcaa	gaggcccgca	ccgatcgccc	ttcccaacag	ttgcgcagcc	3720
tgaaatggcga	atggcgagct	ggctgttttg	gcggatgaga	taagattttc	agctctgatac	3780
agatttaaat	agaaacgaga	agcggctgta	taaaaacagaa	tttgctctggc	ggcagtagcg	3840
cggtggctccc	accttgaccct	atggcgaact	cagaagtga	acgcctgagc	gcccagtggt	3900
gtgtggggctc	ttccctatgcg	agagttaggga	actgcgaggg	atacaataaa	acgaagaggt	3960
cagtcgaaag	actggggcctt	tcgttttatc	tggtgtttgt	cggtgaaacg	ttctcttagat	4020
aggacaaatc	cgccggggagc	ggatttgaac	gtgtcgaaagc	aacgggcccgg	agctgtggcg	4080
gcaggagccg	gccataaac	tgccaggcat	caaattaaagc	agaaaggccat	ccctagcggt	4140
ggccttttttg	gctttctaca	aactcttttt	gtttattttt	ctaaatacat	ttaagtattg	4200
atccgctcat	gagacaataa	ccctgataaa	tgcttcaata	atattgaaaa	aggaagagta	4260
tgagtattca	acattttccgt	gtcgccctta	ttcccttttt	tgccgcat	tttccttctg	4320
tttttgtctc	cccagaaagc	ctgtgtgaag	taaaagatgc	tgaagatcag	ttgggtgcac	4380
gagtggggtta	catcgaaactg	gatctcaaca	gcggtaaagt	ccttgagagt	tttcgcccgc	4440
aagaacgttc	ttcaatgatg	agcactttta	aagttctgct	atgtggcgcg	gtattatccc	4500
gtgtgtgcgc	cgggcaagag	caactcggct	gccgcataca	ctatttctcag	gtactattgg	4560
ttgagtactc	accagtcaca	gaaaagcatc	ttacggatg	catgacagta	agagaattat	4620
gcagtgctgc	cataacacga	agtataaca	cttactctctg	cttactctctg	acaacagtcg	4680
gaggacggga	ggagctaac	gcttttttgc	acaacatggg	ggatcatgta	actcgcttgc	4740
atcgttggaa	accggagctg	aatgaagcca	taccaaacga	cgagcgtgac	accagatgc	4800
ctgtagcatc	ggcaacaacg	ttgcgcaaac	tattaaactg	cgaactactt	actctagctt	4860
cccggcaaca	attaatagac	tggaatggag	gcgataaagt	tgacgagcca	ctgtctcgct	4920
gggccccttc	ggctgcctg	tttatgtctg	ataaatctg	agccggtgag	cggtggctct	4980
gcggtattac	tgccagctgt	gggccagatg	gtaagccctc	cgatctcgt	gttatctaca	5040
gcagcggggag	tcaggcaact	atggatgaac	gaaatagaca	gatcgctgag	ataggtgcct	5100
cactgtattaa	gcattgtgtaa	ctgtcagacc	aagtttactc	atatataact	tagattgatt	5160
taccccggtt	gataatcaga	aaagccccaa	aaacaggaag	atttgtataag	caaatattta	5220
aatgttataac	tgtttaattt	gtttaaattt	cgcgttaaat	tttgttttaa	ctgactactt	5280
ttttaaccaa	tagggcgaaa	tcggcaaaat	cccttataaa	tcaaaagaa	agcccgagat	5340
aggggttgagt	gttgttccag	tttggaacaa	gagtcacaa	ttaaagaacg	ctgactccaa	5400
cgtaacaagg	cgaaaatacgc	ctctatcagg	cgatggccca	ctacgtgaac	catgacccaa	5460
atcaagtttt	ttggggtcga	gggtccgttaa	agcactaaat	cggaacccca	aagggagacc	5520
ccgattttaga	gcttgcaggg	gaaagccggc	gaacgtggcg	agaaaggaag	ggaaggaagc	5580
gaaaggagcg	ggcgctaggg	cgctggcaag	tgtagcggtc	acgctgcgcg	taccaccacc	5640
acccgcgcgc	cttaatgcgc	gcctacaggg	cgcgtaaaag	gatctagggt	aagatccctt	5700
ttgataatct	catgaccaaa	atcccttaac	gtgagttttc	gttccactga	gcctcagacc	5760
ccgtagaaaa	gatcaaaagga	ctctcttgag	atctcttttt	tctgcgcgta	atctgtcgct	5820
gtcaaacaaaa	aaaaccacgc	ctaccagcgg	tggttttgat	gccggctcaa	gagctaccaa	5880
ctcttttttc	gaaggtaact	ggcttcagca	gagcgcagta	accaaatact	gtccttctag	5940

Q95957 Substitute Sequence Listing 0409.TXT

```

tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc 6000
tgctaattcct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg 6060
actcaagacg atagttaccg gataaggcgc agcgggtcgg gtgaacgggg ggttcgtgca 6120
cacagccagc cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat 6180
gagaaagcgc cagcttccc gaagggagaa aggcggacag gtatccggta agcggcaggc 6240
tcggaacagg agagcgacg agggagcttc cagggggaaa cgccctggat ctttatagtc 6300
ctgtcggggt tcgccacct tgacttgagc gtcgattttt gtgatgctcg tcaggggggc 6360
ggagccctat gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgcctgc 6420
cttttgcctca catgttcttt cctgcgttat ccctgattc tgtdggataac cgattaccgc 6480
cctttgagtg agctgtatcc gctgcgcgca gccgaacgac cgagcgcagc gagtcagtga 6540
gcgaggaagc ggaagagcgc ctgatgcggt attttctcct tacgcatctg tgcgggtattt 6600
cacaccgcat atgtgtgact ctacgtacaa tctgctctga tgccgcatag ttaagccagt 6660
atacactccg ctatcgctac gtgactgggt catggctgcy ccccgacacc cgccaacacc 6720
cgctgacgcy ccctgacggg cttgtctgct ccgggcatcc gcttacagac aagctgtgac 6780
cgtctccggg agctgcatgt gtcagagggt ttcaccgtca tcaccgaaac gcgcgagatca 6840
gctcgggtaa agctcatcag cgtggtcgtg cagcgattca catagtctgt cctgtttacc 6900
cgcgctcagc tcggttagtt tctccagaag cgttaatgtc tggcttctga taaagcggcg 6960
catggttaag cggttttttt cctgttttgg cacttgatgc ctccgtgtaa gggggaaattt 7020
ctgttctagg gggttaagtt accgatgaaa cgagagagga gtctcacgat acgggttact 7080
gatgatgaac atgcccgggt actggaacgt tgtgagggtt aacaactggc gggtatggatg 7140
ggcggggacc agagaaaaat cactcagggg caatgccagc gcttcggttaa tacagatgta 7200
gggtgtccact agggtagcca cgagcatcct gcgatgcaga tccggaaact aatgggtgacg 7260
ggcgctgact tccgcgtttc cagactttac gaaacacgga aaccgaagac catctatgtt 7320
gttgctcagg tcgcagacgt ttgacgacg cagtcgcttc acgttcgctc gcgtatcggt 7380
gattcatctt gctaaccagt aaggcaacc cgccagccta gccgggtctc caacgacagg 7440
agcacgatca tgcgcacccg tggccaggac ccaacgctgc ccgaattt 7488

```

<210> 2
 <211> 671
 <212> PRT
 <213> artificial sequence

<220>
 <223> reverse transcriptase

<400> 2
 Thr Leu Asn Ile Glu Asp Glu His Arg Leu His Glu Thr Ser Lys Glu
 1 5 10 15
 Pro Asp Val Ser Leu Gly Ser Thr Trp Leu Ser Asp Phe Pro Gln Ala
 20 25 30
 Trp Ala Glu Thr Gly Gly Met Gly Leu Ala Val Arg Gln Ala Pro Leu
 35 40 45
 Ile Ile Pro Leu Lys Ala Thr Ser Thr Pro Val Ser Ile Lys Gln Tyr
 50 55 60
 Pro Met Ser Gln Glu Ala Arg Leu Gly Ile Lys Pro His Ile Gln Arg
 65 70 75 80
 Leu Leu Asp Ala Gly Ile Leu Val Pro Cys Gln Ser Pro Trp Asn Thr
 85 90 95
 Pro Leu Leu Pro Val Lys Lys Pro Gly Thr Asn Asp Tyr Arg Pro Val
 100 105 110
 Gln Asp Leu Arg Glu Val Asn Lys Arg Val Glu Asp Ile His Pro Thr
 115 120 125
 Val Pro Asn Pro Tyr Asn Leu Leu Ser Gly Leu Pro Pro Ser His Gln
 130 135 140
 Trp Tyr Thr Val Leu Asp Leu Lys Asp Ala Phe Phe Cys Leu Arg Leu
 145 150 155 160
 His Pro Thr Ser Gln Pro Leu Phe Ala Phe Glu Trp Arg Asp Pro Glu
 165 170 175
 Met Gly Ile Ser Gly Gln Leu Thr Trp Thr Arg Leu Pro Gln Gly Phe
 180 185 190
 Lys Asn Ser Pro Thr Leu Phe Asp Glu Ala Leu His Arg Asp Leu Ala
 195 200 205
 Asp Phe Arg Ile Gln His Pro Asp Leu Ile Leu Leu Gln Tyr Val Asp

Q95957 Substitute Sequence Listing 0409.TXT

210	215	220																
Asp	Leu	Leu	Leu	Ala	Ala	Thr	Ser	Glu	Leu	Asp	Cys	Gln	Gln	Gly	Thr			
225	Arg	Ala	Leu	Leu	Gln	Thr	Leu	Gly	Asn	Leu	Gly	Tyr	Arg	Ala	Ser	Ala		
					245					250					255			
Lys	Lys	Ala	Gln	Ile	Cys	Gln	Lys	Gln	Val	Lys	Tyr	Leu	Gly	Tyr	Leu			
			260							265					270			
Leu	Lys	Glu	Gly	Gln	Arg	Trp	Leu	Thr	Glu	Ala	Arg	Lys	Glu	Thr	Val			
		275								280								
Met	Gly	Gln	Pro	Thr	Pro	Lys	Thr	Pro	Arg	Gln	Leu	Arg	Glu	Phe	Leu			
		290				295					300							
Gly	Thr	Ala	Gly	Phe	Cys	Arg	Leu	Trp	Ile	Pro	Gly	Phe	Ala	Glu	Met			
305	Ala	Ala	Pro	Leu	Tyr	Pro	Leu	Thr	Lys	Thr	Gly	Thr	Leu	Phe	Asn	Trp		
					310					315					320			
Gly	Pro	Asp	Gln	Gln	Lys	Ala	Tyr	Gln	Glu	Ile	Lys	Gln	Ala	Leu	Leu			
			340							345					350			
Thr	Ala	Pro	Ala	Leu	Gly	Leu	Pro	Asp	Leu	Thr	Lys	Pro	Phe	Glu	Leu			
		355								360					365			
Phe	Val	Asp	Glu	Lys	Gln	Gly	Tyr	Ala	Lys	Gly	Val	Leu	Thr	Gln	Lys			
		370								375					380			
Leu	Gly	Pro	Trp	Arg	Arg	Pro	Val	Ala	Tyr	Leu	Ser	Lys	Lys	Leu	Asp			
385	Pro	Val	Ala	Ala	Gly	Trp	Pro	Pro	Cys	Leu	Arg	Met	Val	Ala	Ala	Ile		
					405					410					415			
Ala	Val	Leu	Thr	Lys	Asp	Ala	Gly	Lys	Leu	Thr	Met	Gly	Gln	Pro	Leu			
			420							425					430			
Val	Ile	Leu	Ala	Pro	His	Ala	Val	Glu	Ala	Leu	Val	Lys	Gln	Pro	Pro			
		435								440					445			
Asp	Arg	Trp	Leu	Ser	Asn	Ala	Arg	Met	Thr	His	Tyr	Gln	Ala	Leu	Leu			
		450				455					460							
Leu	Asp	Thr	Asp	Arg	Val	Gln	Phe	Gly	Pro	Val	Ala	Ala	Leu	Asn	Pro			
465	Ala	Thr	Leu	Leu	Pro	Pro	Glu	Glu	Gly	Leu	Gln	His	Asn	Cys	Leu			
					485					490					495			
Asp	Ile	Leu	Ala	Glu	Ala	His	Gly	Thr	Arg	Pro	Asp	Leu	Thr	Asp	Gln			
			500							505					510			
Pro	Leu	Pro	Asp	Ala	Asp	His	Thr	Trp	Tyr	Thr	Asn	Gly	Ser	Ser	Leu			
		515								520					525			
Leu	Gln	Glu	Gly	Gln	Arg	Lys	Ala	Gly	Ala	Ala	Val	Thr	Thr	Glu	Thr			
		530				535					540							
Glu	Val	Ile	Trp	Ala	Lys	Ala	Leu	Pro	Ala	Gly	Thr	Ser	Ala	Gln	Arg			
545	Ala	Glu	Leu	Ile	Ala	Leu	Gln	Ala	Leu	Lys	Met	Ala	Glu	Gly	Lys			
					565					570					575			
Lys	Leu	Asn	Val	Tyr	Thr	Asp	Ser	Arg	Tyr	Ala	Phe	Ala	Thr	Ala	His			
			580							585					590			
Ile	His	Gly	Glu	Ile	Tyr	Arg	Arg	Arg	Gly	Leu	Leu	Thr	Ser	Glu	Gly			
		595								600					605			
Lys	Glu	Ile	Lys	Asn	Lys	Asp	Glu	Ile	Leu	Ala	Leu	Leu	Lys	Ala	Leu			
		610				615					620							
Phe	Leu	Pro	Lys	Arg	Leu	Ser	Ile	Ile	His	Cys	Pro	Gly	His	Gln	Lys			
625	Gly	His	Ser	Ala	Glu	Ala	Arg	Gly	Asn	Arg	Met	Ala	Asp	Gln	Ala			
					645					650					655			
Arg	Lys	Ala	Ala	Ile	Thr	Glu	Thr	Pro	Asp	Thr	Ser	Thr	Leu	Leu				
			660							665				670				

<210> 3
 <211> 7488
 <212> DNA
 <213> artificial sequence

Q95957 Substitute Sequence Listing 0409.TXT

<220>

<223> plasmid DNA

<400> 3

ccgacccat	cgaatgtgtc	aaaacctttc	cgcgatatgc	atgatagcgc	ccggaagaga	60
gtcaattcag	gggtgtgaat	gtgaaaccag	taacgttata	cgatgtgcga	gagtatgccg	120
gtgtctctta	tcagaccgtt	tcctcgctgg	tgaaccaggc	cagccacggt	tctgcgaaaa	180
cgcgggaaaa	agtggaagcg	gcgatggcgg	agctgaatta	cattcccaac	cgctgggcac	240
aacaactggc	gggcaaacag	tcgttgctga	ttggcgctgc	cacctccagt	ctggccctgc	300
acgcgcctgc	gcgaattgtc	gcggcgatta	aatctcgctc	cgatcaactg	gggtgccacg	360
tggtgtgtgc	gatggtagaa	cgaaagcgcg	tcgaagctgc	taaacggcgc	gtgcacaatc	420
tctctcgcca	acgcgtcagt	gggctgatca	ttaactatcc	gctggatgac	caggatgcca	480
ttgctgtgga	agctgcctgc	actaatgttc	cggcgttatt	tcttgatgtc	tctgaccaga	540
cacccatcaa	cagtattatt	ttctcccatg	aagacggtac	gcgactgggc	gtggagcatc	600
tggtcgattt	gggtcaccag	caaatcgctc	tgttagcggg	cccattaaat	tctgtctcgg	660
cgctctcgcg	tctggctggc	tggtcataat	atctcactcg	caatcaaat	cagccgatag	720
cggaaccgga	aggcgactgg	agtgccatgt	ccggttttca	acaaaccatg	caaatgtctga	780
atgaggcgcat	cgctcccatc	gcgatgtctg	ttgccaacga	tcagatggcg	ctgggcccga	840
tgccgcgcct	taccgagctc	gggtctgcgc	ttggtgcgga	tatctcgcta	gtgggatacg	900
acgataccga	acgataccga	gtttatatcc	cgccgttaac	caccatcaaa	caggattttc	960
gcctgtggtg	gcgaaccagc	gtggaccgct	tgctgcaact	ctctcagggc	caggcggtga	1020
agggcaatca	gctgtgtccc	gtctcactgg	tgaaaagaaa	aaccacccgt	cgccccaata	1080
cgcaaacctgc	ctctccccgc	cgctgtggcg	atctcataat	cgctcggtga	cagcaggttt	1140
cccgaactgc	aagcgggcag	tgagcgcac	gcaattaatg	tgagttagct	cactcattag	1200
gcacaattct	catgtttgac	agcttatcat	cgactgcacg	gtgaccaaat	gctcttggtg	1260
tcaggagctc	atcgggaagt	gtggatggc	tgctcaggct	gtcaatacact	gcttaattct	1320
tgctgctcaa	ggcgactccc	cgttctggat	aatgtttttt	cgcccgagat	cataacgggt	1380
ctggcaataa	ttctgaaatg	agctgttgac	aattaatcat	cgcctgtgat	aatgtgtgga	1440
attgtgtagcg	gatacaaatc	tgaaattcta	agattttgtga	ggggataaaca	atttcacaca	1500
ggaaacagaa	tatgacctta	aatatagaag	atgagcatcg	gctacatgag	acctcaaaag	1560
agccagatgt	ttctctaggg	ttccactggc	tgcttgattt	tcctcaggct	tgggcggaac	1620
ccggggcgat	gggactggca	gtttccgca	ctctcttgat	catacctctg	aaagcaacat	1680
ctacccccgt	gtccataaaa	caatacccca	tgtcacaaag	agccagactg	gggatcaaac	1740
ccccacacga	gagactgttg	gaacaacggga	tcctgtgtac	ctgccagctc	ctctggaaac	1800
gcgccctctg	accggttaa	aaaccaggga	ctaattgata	taggcctctg	caggatctga	1860
gagaagtcaa	caagcgggtg	gaagacatcc	acccacccgt	gcccaaccct	tacaacctct	1920
tgagcgggtg	cccaccgtcc	caccagtgg	acactgtgct	tgatttaag	gatgcctttt	1980
tctgtctgag	actccaccct	accagtacg	ctctcttcgc	ctttgagtgg	agagatcttg	2040
agatgggaat	ctcaggacaa	ttgacctgga	ccgactccc	acaggggttc	aaaaaacagt	2100
ccaccctgtt	tgatgagcga	gtgcacagag	acctagcaga	cttccggatc	cagcacccag	2160
acttgatcct	gctacagtac	gtggatgat	tactgtggc	cgccacttct	gagctagact	2220
gccacaagag	tactcgggcc	ctgttacaac	acctagggaa	cctcggtgat	cgctgcctcg	2280
ccaagaaagc	caaaatttgc	cagaaacagg	tcaagtatct	ggggatctct	ctaaagagg	2340
gtcagagatg	gctgactgag	gcgagaaaag	agactgtgat	ggggcagcct	actccgaaga	2400
ccccctgcaga	actaaggagg	ttctctagga	cggcaggctt	ctgtcgctct	tggtatccctg	2460
ggttttgcaga	aatggcagcc	ccctgttacc	ctctcaccaa	aacggggagt	ctgtttaatt	2520
ggggcgccaga	cccaacaaa	gcctatcaag	aaatcaagca	agctcttcta	actgccccag	2580
ccctgggtgt	gcagagattg	actaaagcct	ttgaactctt	tgctcgacga	aagcagggct	2640
acgccaaaag	tgctctaacg	caaaaactgg	gaccttggcg	tcggccgggt	gcttacctgt	2700
ccaaaagcct	agaccagatg	gcagctgggt	ggcccccttg	ctcagcgatg	ctgcagacca	2760
ttcccgatct	gacaaagatg	cgaggcaagc	taaccatggg	acagccacta	gtcattcttg	2820
ctccccctgc	agtagagcca	ctagtcaaac	aaacccccga	ccgctgtgct	ctcaacgcc	2880
ggatgactca	ctatcagccc	ttgcttttgg	acacggaccg	ggctcagttc	ggaccggtgg	2940
tagccttgaa	cccggctacg	gtctcccac	tgcttaggga	agggctgcga	cgaactgcc	3000
ttgatattct	ggccgaagcc	cacggaaccc	gacctgacct	aacggagacc	cgctctccag	3060
acgcgcacca	caactgtgat	acgaatggaa	gcagttcttt	acaagaggga	cagcgtaaag	3120
cgggagctgc	ggtagaccac	gagaccgagg	taactctggc	taaaagcctg	ccagccggga	3180
catctcgctca	cggggctgaa	ctgatagcac	tcacccagg	cctaaagatg	ctgaaggata	3240
agaagctcaaa	tgattatact	gatagcgctt	tgcttttgc	tactgcccat	atccatggag	3300
aaatatcacg	aaggcgtggg	ttgctcacat	cagaaggcaa	agagatacaa	aataaagacg	3360
agatcttggc	ctctactaaa	gcctcttttc	tgccccaaag	acttagcata	atccattgtc	3420
caggacatca	aaaggggac	agcgcggagg	ctagaggcaa	ccggaatggc	gaccagcgc	3480
ccgaaaggc	agccatcaca	gagactccag	acacctctac	ctctctccat	caccatcac	3540

Q95957 Substitute Sequence Listing 0409.TXT

atcactagtc	tagagtcgac	ctcgaggcaa	gcttgccact	ggccgtcggt	ttacaacgct	3600
gtgactggga	aaaccctggc	gttaccacaac	ttaactcgct	tgcagcacat	ccccctttcg	3660
ccagctggcg	taatagcgaa	gaagcccgca	ccgatcgccc	ttcccaacag	ttgcgcagcc	3720
tgaatggcga	atggcagcgt	ggctgttttt	gcggatgaga	taagattttc	agcctgatac	3780
agattaaatc	agaacgcaga	agcggctgta	taaaaacagaa	tttgctcgcc	ggcagtagcg	3840
cggtgggtccc	accctgacccc	atgccgaact	cagaagtga	acgcccgtag	gccgatggta	3900
gtgtgggtggt	tcccatcgcg	agagtagggga	actgccagcc	atcaaatata	acgaaggcct	3960
cagtcgaaag	actgggcttt	tcgtttttatc	tggtgtttgt	cggtgaacgc	tctcctgagt	4020
aggacaatac	gcccgggagc	ggattttgaac	gttgcgaaagc	aacggcccgg	agggtggcgg	4080
cgaggacgcc	gcccataaac	tgccaggcat	caaatataagc	agaaggccat	cctgacggat	4140
ggcctttttg	cgttttctaca	aactcttttt	gtttattttt	ctaaatacat	tcaaatatgt	4200
atccgctcat	gagacaataa	ccctgataaa	tgcttcaata	atatgaaaa	aggaagagta	4260
tgagtattca	acatttcctg	gtcgccctta	tccctttttt	tcggtcattt	gtccttccctg	4320
tttttgctca	ccagaaacgc	ctggtgaaag	taaaagatgc	tgaagatcag	tgggtgcac	4380
gagtggggta	catcgaactg	gatctcaaca	gcggttaagat	ccttgagagt	tttcgccccg	4440
aagaacgttc	tccaatgatg	agcactttta	aagtctctgt	attgtggcgcg	gtattattccc	4500
gtgtttgacg	cgggcaagtg	caactcggct	gccgcataca	ctattctcat	acaactatgg	4560
tttgagtactc	accagtcaac	gaaagacatc	ttacggatgc	catgacagta	agagaattat	4620
gcagtgctgc	cataacctag	agtataaca	ctgcggccaa	cttactctc	acaacgatcg	4680
gaggacgcga	ggagctaaac	gcttttttgc	acaaactggg	ggatcatgta	actcgcttgc	4740
atcgttggga	accggagctg	ataagaagca	taccaaacga	cgagcgtgac	accacgatgc	4800
ctgtagcaatc	ggcaacaacg	ttgcgcaaac	tattaaactgc	cgaaactact	actctagctt	4860
ccgggcaaca	ttcaatagag	tggatggagg	cggataaagt	ttcagaacac	cttcctcgct	4920
cgggcccttcc	ggctggctgg	tttatgtctg	ataaatctgg	agccggtgag	cggtgggtctc	4980
gcggtatcat	tgcagcactg	gggcccagat	gtaagccctc	cgatctgta	gttatctaca	5040
cgacggggag	tcagggcaact	ttggatgaac	gaaatagaca	gatcgctgag	ataggtgcct	5100
catctgattaa	gcatttggtaa	ctgtcgagac	aagttttact	atatatact	tagattgatt	5160
taccccgggt	gataaactaga	aaagccccaa	aaacagggaag	attgtataag	caaatattta	5220
aattgtaaac	gttaataatt	tgtaaaaaat	ccgtttaaatt	ttttgttaaa	tcagctctct	5280
ttttaaccaa	tagggccgaa	tcggcaaaat	cccttataaa	tcaaaaagaa	accgccagat	5340
aggggtgagt	gttgttccag	tttggaaaca	gagtcacta	ttaaaagacg	tggactccaa	5400
cgttaacggt	cgaaaaacgc	tctatcaggg	cgatggccca	tacagtgaac	catcaccaaa	5460
atcaaatgttt	ttggggctga	gggtccgttaa	agcactaaat	cgaaacccca	aagggagccc	5520
ccgattttaga	gctttgacgg	gaacgctggc	gaacgctggc	agaaagggaag	gggaagaaagc	5580
gaaaggagcg	ggcgctaggg	cgctggcaag	tgtagcggct	acgctgcgcg	taaccaccac	5640
accgcgcgcg	cttaaatgcg	cgctacaggg	cgcgtaaaag	gatctagggt	aagatccctt	5700
ttgataatct	catgaccaaa	atcccttaac	gtgagttttc	gttccactga	gcgtcgagcc	5760
ccgtagaaaa	gatcaaaagg	tcttcttgag	atcctttttt	tctcgcgcta	attctgtgct	5820
tgcaaaaaca	aaaaccacgc	ctaccagcgg	tggtttgttt	gccggatcaa	gagctaccaa	5880
ctcttttttc	gaaggtaact	ggccttcagca	gagcgcagat	accaaatact	gtccttctag	5940
tgtatccgta	gttaggccac	cacttcaaga	actctgtagc	accgcctaca	taccctgcgt	6000
gtctaactct	gttaccagtg	gctcgtgcca	gtggcgataa	gtcgtgtctt	gggtccgttg	6060
actcaagacg	atagttaccg	gataaggcgc	agcggctcgg	ctgaacgggg	gggtcgctga	6120
cacagcccgag	ctttggagcga	acgaactaca	cgaaactgag	ataactacag	cgtagactga	6180
gagaagcgcg	cacgcttccc	gaaggggagaa	agggcgacag	gtatccggta	agcggcaggg	6240
tcggaacagc	agagcgcagc	agggagcttc	aggggggaaa	cgctgtgtat	ctttatagtc	6300
ctgtcgggttt	tcgccacctc	tgacttgagc	gtcgattttt	gtgatgctcg	tcagggggggc	6360
ggagccttat	gaaaaaacgc	agcaacgcgg	ctttttttacg	gttctgggcc	ttttgtggc	6420
cttttgctca	catgtttctt	ctctgcttat	ccccgtatc	tgtggataac	cgattaccgc	6480
cttttgatga	agctgatacc	gctcgcgcga	gccgaacgac	cgagcgcagc	gagtcagtga	6540
cgagggaagc	ggaagagcgc	ctgatgcggg	attttctcct	tacgcatctg	tgcggtattt	6600
cacaccgcgat	attggtgcact	ctcagtacaa	tctgctctga	tggcgcatag	ttaagccagt	6660
atacatctcg	ctctcgctac	gtgactgggt	catggctgcy	ccccgacacc	cgcccaacc	6720
cgctgacgcg	ccctgacggg	cttgtctgct	ccgggcatcc	gcttacagac	aagctgtgac	6780
cgctctcggg	agctgcattg	gtcagagggt	ttcacggtca	taccggaac	cgccgaggca	6840
gctcgcggtaa	agctcatcag	gtgggtcgtg	cagcgattca	cagatgtctg	gcttctcatc	6900
cgcgctcagc	tcgttgagtt	tcttcagaag	cgttaatgtc	tggcttctga	taaaagcggc	6960
catgtttaag	ggggtttttt	ccgttttggg	cacttgatgc	ctccgtgttaa	gggggaattt	7020
ctgttcatgg	gggtaatgat	accgatgaaa	cgagagagga	tgctcacgat	accgggttact	7080
gatgatgaac	atgcccgggt	actggaacgt	tgtgagggta	aacaactggc	ggatgggatg	7140
gggggggacc	agagaaaaat	cactcagggt	caatgacagc	gcttctgttaa	tacagatgta	7200
gggtgtccac	ccggtatgcca	cgacatctct	cgcatgcaga	tccggaacat	aattggtcgag	7260
ggcgtcgact	tcggcgtttc	cagactttac	gaaacacgga	aaccgaagac	catctcatgt	7320

Q95957 Substitute Sequence Listing 0409.TXT

gttgctcagg tcgcagacgt ttgcagcag cagtcgcttc acgttcgctc gcgtatcggt
 gattcattct gctaaccagt aaggcaaccc cgccagccta gccgggtcct caacgacagg
 agcacgatca tgcgcaccgc tggccaggac ccaacgctgc ccgaattt

7380
 7440
 7488

<210> 4
 <211> 671
 <212> PRT
 <213> artificial sequence

<220>
 <223> reverse transcriptase

<400> 4

Thr Leu Asn Ile Glu Asp Glu His Arg Leu His Glu Thr Ser Lys Glu
 1 5 10 15
 Pro Asp Val Ser Leu Gly Ser Thr Trp Leu Ser Asp Phe Pro Gln Ala
 20 25 30
 Trp Ala Glu Thr Gly Gly Met Gly Leu Ala Val Arg Gln Ala Pro Leu
 35 40 45
 Ile Ile Pro Leu Lys Ala Thr Ser Thr Pro Val Ser Ile Lys Gln Tyr
 50 55 60
 Pro Met Ser Gln Glu Ala Arg Leu Gly Ile Lys Pro His Ile Gln Arg
 65 70 75 80
 Leu Leu Asp Asn Gly Ile Leu Val Pro Cys Gln Ser Pro Trp Asn Thr
 85 90 95
 Pro Leu Leu Pro Val Lys Lys Pro Gly Thr Asn Asp Tyr Arg Pro Val
 100 105 110
 Gln Asp Leu Arg Glu Val Asn Lys Arg Val Glu Asp Ile His Pro Thr
 115 120 125
 Val Pro Asn Pro Tyr Asn Leu Leu Ser Gly Leu Pro Pro Ser His Gln
 130 135 140
 Trp Tyr Thr Val Leu Asp Leu Lys Asp Ala Phe Phe Cys Leu Arg Leu
 145 150 155 160
 His Pro Thr Ser Gln Pro Leu Phe Ala Phe Glu Trp Arg Asp Pro Glu
 165 170 175
 Met Gly Ile Ser Gly Gln Leu Thr Trp Thr Arg Leu Pro Gln Gly Phe
 180 185 190
 Lys Asn Ser Pro Thr Leu Phe Asp Glu Ala Leu His Arg Asp Leu Ala
 195 200 205
 Asp Phe Arg Ile Gln His Pro Asp Leu Ile Leu Leu Gln Tyr Val Asp
 210 215 220
 Asp Leu Leu Leu Ala Ala Thr Ser Glu Leu Asp Cys Gln Gln Gly Thr
 225 230 235 240
 Arg Ala Leu Leu Gln Thr Leu Gly Asn Leu Gly Tyr Arg Ala Ser Ala
 245 250 255
 Lys Lys Ala Gln Ile Cys Gln Lys Gln Val Lys Tyr Leu Gly Tyr Leu
 260 265 270
 Leu Lys Glu Gly Gln Arg Trp Leu Thr Glu Ala Arg Lys Glu Thr Val
 275 280 285
 Met Gly Gln Pro Thr Pro Lys Thr Pro Arg Gln Leu Arg Glu Phe Leu
 290 295 300
 Gly Thr Ala Gly Phe Cys Arg Leu Trp Ile Pro Gly Phe Ala Glu Met
 305 310 315 320
 Ala Ala Pro Leu Tyr Pro Leu Thr Lys Thr Gly Thr Leu Phe Asn Trp
 325 330 335
 Gly Pro Asp Gln Gln Lys Ala Tyr Gln Glu Ile Lys Gln Ala Leu Leu
 340 345 350
 Thr Ala Pro Ala Leu Gly Leu Pro Asp Leu Thr Lys Pro Phe Glu Leu
 355 360 365
 Phe Val Asp Glu Lys Gln Gly Tyr Ala Lys Gly Val Leu Thr Gln Lys
 370 375 380
 Leu Gly Pro Trp Arg Arg Pro Val Ala Tyr Leu Ser Lys Lys Leu Asp
 385 390 395
 Page 7

Q95957 Substitute Sequence Listing 0409.TXT

```

385          390          395          400
Pro Val Ala Ala Gly Trp Pro Pro Cys Leu Arg Met Val Ala Ala Ile
              405              410              415
Ala Val Leu Thr Lys Asp Ala Gly Lys Leu Thr Met Gly Gln Pro Leu
              420              425              430
Val Ile Leu Ala Pro His Ala Val Glu Ala Leu Val Lys Gln Pro Pro
              435              440              445
Asp Arg Trp Leu Ser Asn Ala Arg Met Thr His Tyr Gln Ala Leu Leu
              450              455              460
Leu Asp Thr Asp Arg Val Gln Phe Gly Pro Val Val Ala Leu Asn Pro
              465              470              475
Ala Thr Leu Leu Pro Leu Pro Glu Glu Gly Leu Gln His Asn Cys Leu
              485              490              495
Asp Ile Leu Ala Glu Ala His Gly Thr Arg Pro Asp Leu Thr Asp Gln
              500              505              510
Pro Leu Pro Asp Ala Asp His Thr Trp Tyr Thr Asn Gly Ser Ser Leu
              515              520              525
Leu Gln Glu Gly Gln Arg Lys Ala Gly Ala Ala Val Thr Thr Glu Thr
              530              535              540
Glu Val Ile Trp Ala Lys Ala Leu Pro Ala Gly Thr Ser Ala Gln Arg
              545              550              555
Ala Glu Leu Ile Ala Leu Thr Gln Ala Leu Lys Met Ala Glu Gly Lys
              565              570              575
Lys Leu Asn Val Tyr Thr Asp Ser Arg Tyr Ala Phe Ala Thr Ala His
              580              585              590
Ile His Gly Glu Ile Tyr Arg Arg Arg Gly Leu Leu Thr Ser Glu Gly
              595              600              605
Lys Glu Ile Lys Asn Lys Asp Glu Ile Leu Ala Leu Leu Lys Ala Leu
              610              615              620
Phe Leu Pro Lys Arg Leu Ser Ile Ile His Cys Pro Gly His Gln Lys
              625              630              635
Gly His Ser Ala Glu Ala Arg Gly Asn Arg Met Ala Asp Gln Ala Ala
              645              650              655
Arg Lys Ala Ala Ile Thr Glu Thr Pro Asp Thr Ser Thr Leu Leu
              660              665              670

```

<210> 5
 <211> 17
 <212> DNA
 <213> artificial sequence

<220>
 <223> primer

<400> 5
 gtggaattgt gagccga

17

<210> 6
 <211> 29
 <212> DNA
 <213> artificial sequence

<220>
 <223> primer

<400> 6
 cggaattccc gcgtccaaca gtctctgta

29

<210> 7
 <211> 17
 <212> DNA
 <213> artificial sequence

Q95957 Substitute Sequence Listing 0409.TXT

<220>
<223> primer

<400> 7
tgggagtctg gtccagg

17

<210> 8
<211> 24
<212> DNA
<213> artificial sequence

<220>
<223> primer

<400> 8
cggaattctg gtaccctgcc agtc

24

<210> 9
<211> 671
<212> PRT
<213> artificial sequence

<220>
<223> reverse transcriptase

<400> 9
Thr Leu Asn Ile Glu Asp Glu His Arg Leu His Glu Thr Ser Lys Glu
1 5 10 15
Pro Asp Val Ser Leu Gly Ser Thr Trp Leu Ser Asp Phe Pro Gln Ala
20 25 30
Trp Ala Glu Thr Gly Gly Met Gly Leu Ala Val Arg Gln Ala Pro Leu
35 40 45
Ile Ile Pro Leu Lys Ala Thr Ser Thr Pro Val Ser Ile Lys Gln Tyr
50 55 60
Pro Met Ser Gln Glu Ala Arg Leu Gly Ile Lys Pro His Ile Gln Arg
65 70 75 80
Leu Leu Asp Gln Gly Ile Leu Val Pro Cys Gln Ser Pro Trp Asn Thr
85 90 95
Pro Leu Leu Pro Val Lys Lys Pro Gly Thr Asn Asp Tyr Arg Pro Val
100 105 110
Gln Asp Leu Arg Glu Val Asn Lys Arg Val Glu Asp Ile His Pro Thr
115 120 125
Val Pro Asn Pro Tyr Asn Leu Leu Ser Gly Leu Pro Pro Ser His Gln
130 135 140
Trp Tyr Thr Val Leu Asp Leu Lys Asp Ala Phe Phe Cys Leu Arg Leu
145 150 155 160
His Pro Thr Ser Gln Pro Leu Phe Ala Phe Glu Trp Arg Asp Pro Glu
165 170 175
Met Gly Ile Ser Gly Gln Leu Thr Trp Thr Arg Leu Pro Gln Gly Phe
180 185 190
Lys Asn Ser Pro Thr Leu Phe Asp Glu Ala Leu His Arg Asp Leu Ala
195 200 205
Asp Phe Arg Ile Gln His Pro Asp Leu Ile Leu Leu Gln Tyr Val Asp
210 215 220
Asp Leu Leu Leu Ala Ala Thr Ser Glu Leu Asp Cys Gln Gln Gly Thr
225 230 235 240
Arg Ala Leu Leu Gln Thr Leu Gly Asn Leu Gly Tyr Arg Ala Ser Ala
245 250 255
Lys Lys Ala Gln Ile Cys Gln Lys Gln Val Lys Tyr Leu Gly Tyr Leu
260 265 270
Leu Lys Glu Gly Gln Arg Trp Leu Thr Glu Ala Arg Lys Glu Thr Val
275 280 285
Met Gly Gln Pro Thr Pro Lys Thr Pro Arg Gln Leu Arg Glu Phe Leu
Page 9

095957 Substitute Sequence Listing 0409.TXT

Gly 305 Ala	290 Thr Ala Gly Phe Cys 310				295 Arg Leu Trp Ile Thr 330				300 Gly Phe Ala Glu Met 320				
	Ala	Pro	Leu	Thr 325	Pro	Leu	Thr	Lys	Thr 330	Gly	Thr	Leu	Phe 335
Gly	Pro	Asp	Gln 340	Gln	Lys	Ala	Tyr	Gln 345	Ile	Lys	Gln	Ala 350	Leu
Thr	Ala	Pro 355	Ala	Leu	Gly	Leu	Pro 360	Asp	Leu	Thr	Lys	Pro 365	Phe
Phe	Val 370	Asp	Glu	Lys	Gln	Gly 375	Tyr	Ala	Lys	Gly	Val 380	Leu	Thr
Leu 385	Gly	Pro	Trp	Arg	Arg 390	Pro	Val	Ala	Tyr	Leu 395	Ser	Lys	Lys
Pro	Val	Ala	Ala	Gly 405	Trp	Pro	Pro	Cys	Leu 410	Arg	Met	Val	Ala
Ala	Val	Leu	Thr 420	Lys	Asp	Ala	Gly	Lys 425	Leu	Thr	Met	Gly	Gln
Val	Ile	Leu 435	Ala	Pro	His	Ala	Val 440	Glu	Ala	Leu	Val	Lys 445	Gln
Asp	Arg 450	Trp	Leu	Ser	Asn	Ala 455	Arg	Met	Thr	His	Tyr 460	Gln	Ala
Leu 465	Asp	Thr	Asp	Arg	Val 470	Gln	Phe	Gly	Pro	Val 475	Val	Ala	Leu
Ala	Thr	Leu	Leu	Pro 485	Leu	Pro	Glu	Glu	Gly 490	Leu	Gln	His	Asn
Asp	Ile	Leu	Ala 500	Glu	Ala	His	Gly	Thr 505	Arg	Pro	Asp	Leu	Thr
Pro	Leu	Pro 515	Asp	Ala	Asp	His	Thr 520	Trp	Tyr	Thr	Asp	Gly 525	Ser
Leu	Gln 530	Glu	Gly	Gln	Arg	Lys 535	Ala	Gly	Ala	Ala	Val 540	Thr	Thr
Glu 545	Val	Ile	Trp	Ala	Lys 550	Ala	Leu	Pro	Ala	Gly 555	Thr	Ser	Ala
Ala	Glu	Leu	Ile	Ala 565	Leu	Thr	Gln	Ala	Leu 570	Lys	Met	Ala	Glu
Lys	Leu	Asn	Val 580	Tyr	Thr	Asp	Ser	Arg 585	Tyr	Ala	Phe	Ala	Thr
Ile	His	Gly 595	Glu	Ile	Tyr	Arg	Arg 600	Arg	Gly	Leu	Leu	Thr 605	Ser
Lys	Glu 610	Ile	Lys	Asn	Lys	Asp 615	Ser	Ile	Leu	Ala	Leu 620	Lys	Ala
Phe 625	Leu	Pro	Lys	Arg	Leu 630	Ser	Ile	Ile	His	Cys 635	Pro	Gly	His
Gly	His	Ser	Ala	Glu 645	Ala	Arg	Gly	Asn	Arg 650	Met	Ala	Asp	Gln
Arg	Lys	Ala	Ala 660	Ile	Thr	Glu	Thr	Pro 665	Asp	Thr	Ser	Thr	Leu